Canadian Studies in Population (2024) 51:2
https://doi.org/10.1007/s42650-024-00080-6

RESEARCH NOTE

Models for Estimating Intrinsic r and the Mean Age of a Population at Stability: Evaluations at the National and Sub-national Level

David A. Swanson[1,2,3]

Received: 17 September 2023 / Accepted: 25 February 2024 / Published online: 6 March 2024

Abstract
Using Canada's provinces and territories in conjunction with the "Cohort Change Ratio" approach to generating a stable population, I test the accuracy of two regression models constructed from national-level data designed to estimate two factors of a population at stability from initial conditions at the sub-national levels: (1) its constant rate of change, denoted here by ***r'***; and (2) mean population age. In a test of accuracy at the national level I find that these models provide reasonably accurate estimates. In the tests at the subnational level, the accuracy, as expected, is less, but the results indicate that the national level models provide estimates that are useful. The models are useful because they are tractable and provide information not available from the traditional analytical approaches. Evaluating these models also provides the opportunity to look at Canada's provinces and territories from a stable population perspective. The findings support the use of: (1) The Cohort Change Ratio approach in examining stable population concepts; and (2) the two regression models for estimating ***r'*** and the mean age of a population at stability. They also show that there are connections between initial conditions and stability that have been overlooked. This knowledge gap may be due to the fact that widespread knowledge and acceptance of the ergodic nature of the "age structure factor," have served to mask the possibility that ergodicity does not always apply to other factors. Further exploration of these potential linkages appears to be in order.

Keywords Canada · Cohort Change Ratio · Data Modeling Culture · Demographic Dynamics · Ergodicity · Numerical Analysis · Paridigm Shift · Provinces · Regression · Stable Population Theory · Territories

✉ David A. Swanson
dswanson@ucr.edu

[1] Department of Sociology, University of California Riverside, Riverside, CA 92521, USA

[2] Center for Studies in Demography & Ecology, University of Washington, Seattle, WA 98105, USA

[3] Population Research Center, Portland State University, Portland, OR 97201, USA

Résumé

En m'appuyant sur les données des territoires et des provinces du Canada, et en utilisant la méthode CCR (« Cohort Change Ratio») visant à générer une population stable, j'ai testé l'exactitude de deux modèles de régression élaborés à partir de données nationales et conçus pour évaluer deux composantes d'une population stable à partir des conditions initiales et à des échelles infranationales, nommément: 1) le taux de changement constant, représenté par « r'», et 2) l'âge moyen. En effectuant un test d'exactitude à l'échelle nationale, j'ai constaté que les estimations fournies par ces deux modèles étaient assez précises. Pour les tests à l'échelle infranationale en revanche, la précision s'avérait moindre comme prévu, mais les résultats suggèrent que les modèles s'appuyant sur des données nationales offrent des estimations utiles, notamment parce qu'ils sont faciles à mettre en œuvre et qu'ils fournissent des renseignements que les approches analytiques traditionnelles ne permettent pas d'obtenir. L'évaluation de ces modèles permet également d'étudier la population des provinces et des territoires du Canada sous l'angle de la « stabilité». Les résultats justifient donc l'utilisation: 1) de la méthode CCR pour étudier le concept de « stabilité» de la population, et 2) des deux modèles de régression pour déterminer le taux de changement constant « r'» ainsi que l'âge moyen d'une population stable. Ces résultats montrent également que certains liens existants entre les conditions initiales et la stabilité ont été ignorés, et que ces lacunes pourraient s'expliquer par le fait que la connaissance ainsi que l'acceptation étendues de la nature ergodique de la « structure des âges» ont concouru à masquer la possibilité qu'une telle ergodicité ne s'applique pas toujours à d'autres facteurs. Il est donc nécessaire d'explorer plus avant ces liens potentiels.

1 Introduction

Stable population theory is well-established and widely used (Carmichael, 2016: 343–351; Caswell, 2001; Coale, 1972; Coale & Demeny, 1966; Dublin & Lotka, 1925; Lopez, 1961; Lotka, 1907; Popoff & Judson, 2004; Preston et al., 2001; Sharpe & Lotka, 1911; United Nations, 1968, 2002; Yusuf et al., 2014: 279–301). Inextricably linked to stable population theory are: (1) its constant rate of population change, known as "***intrinsic r***" (Caswell, 2001; Coale, 1972; Coale & Demeny, 1966; Dublin & Lotka, 1925; Lotka, 1907; Preston et al., 2001; Sharpe & Lotka, 1911; United Nations, 1968); and (2) the concept of ergodicity, whereby at stability a population has "forgotten" its initial age structure (Arthur, 1982; Caswell, 2001; Cohen, 1979; Stott et al., 2010; Tuljapurkar, 1982).

As demonstrated by Swanson et al. (2016), Cohort Change Ratios (CCRs) provide a valid, useful, and tractable means of examining the concept of a stable population. This approach takes into account all genders as well as migration, which extends the traditional definition of the constant rate of change in a stable population, ***intrinsic r*** (or more simply, ***r***), which typically was applied to females only and was based on the idea of a population closed to migration. That is, ***r*** is perceived as accommodating only two components of population change, births and deaths, and not the third component, migration. However, as shown by Swanson et al. (2016), the CCR approach easily accommodates not only births and deaths, but also migration and all genders. Given this, this paper uses ***r'*** to designate the constant rate of population change in a stable population generated using the CCR approach.

Swanson et al. (2016) introduced the Index of Stability (***S***) as a measure to determine if a population is stable. They also used this index (along with other predictor variables representing initial conditions) in multiple regression models based on 62 national populations to estimate the number of years before the population in question becomes quasi-stable and ***r'***. They found that their models performed reasonably well.

These results suggest that while ergodicity applies to a population's initial age structure, it may not universally apply to other initial conditions. This possibility is supported by Stott et al. (2010: 242) who find that "reducible" models (those in which parameterized transition rates do not facilitate pathways from all stages to all other stages) based on population projection matrices are sometimes ergodic but may be non-ergodic (the model exhibits two or more stable asymptotic states with different asymptotic stable growth rates, which depend on the initial stage structure used in the population projection). These findings imply that the widespread knowledge and acceptance of ergodicity in regard to the "age structure factor," both in its strong and weak forms (Arthur, 1982), may be "masking" the possibility that ergodicity does not always apply to other factors.

I add to this work by showing that two multiple regression models, each of which is based on the same 62 national populations used by Swanson et al. (2016), perform reasonably well in estimating, respectively: (1) the constant rate of population change upon attaining stability, ***r'*** from its initial mean age and the initial rate of population change (**r**), and (2) the mean age of a population upon attaining stability from its initial mean age and the initial rate of population change, **r**.

"***r'***" is of interest because, like its more limited counterpart, ***r***, it is a defining characteristic of a stable population (Caswell, 2001; Coale, 1972; Coale & Demeny, 1966; Popoff & Judson, 2004; Preston et al., 2001; United Nations, 1968, 2002). Mean age at stability is of interest because it is widely used as a summary measure of population aging (Preston et al., 1989; United Nations, 2017). Unlike Swanson et al. (2016), I divide the 62 countries into two randomly selected groups: (1) a 30-case "training" set, which I use to build the regression model; and (2) a 32 case "validation" set, which I use to evaluate the models. I then explore the extension of these two models to the sub-national level by testing their accuracy on the populations of Canada's provinces and territories and then discussing the results.

The reminder of this paper is composed of five sections. In the next one (II), I briefly discuss the CCR method. In Section III, I describe the approach to the concept of a stable population. Section IV introduces the subnational data (Canada's provinces and territories) and describes the two regression models they are used to evaluate. Section V provides the results and Section VI discusses them.

2 Cohort Change Ratios

The Cohort Change Ratio (CCR) method moves a population by age (and sex) from time t to time t + k using CCRs computed from data in the two most recent censuses (Swanson et al. 2016). It consists of two steps. The first uses existing data to develop CCRs and the second applies the CCRs to the cohorts of the launch year population to move them into the future. The second step can be repeated infinitely, with the projected population serving as the launch population for the next projection cycle. The formula for the first step, the development of a CCR is:

$$nCCR_{x,i} = nP_{x,i,t}/nP_{x-k,i,t-k} \tag{1}$$

where

${}_{n}P_{x,i,t}$ is the population aged x to x + n in area i at the most recent census (t),

${}_{n}P_{x\text{-}k,i,t\text{-}k}$ is the population aged x-k to x-k + n in area i at the 2nd most recent census (t-k),

k is the number of years between the most recent censuses at time t for area i and the one preceding it for area i at time t-k.

The basic formula for the second step, moving the cohorts of a population into the future is:

$$nP_{x+k,i,t+k} = \left(nCCR_{x,i}\right) * \left(nP_{x,i,t}\right) \tag{2}$$

where

${}_{n}P_{x+k,i,t+k}$ is the population aged x + k to x + k + n in area i at time t + k.

${}_{n}CCR_{x,i}$ ${}_{n}P_{x,i,t}$ / ${}_{n}P_{x\text{-}k,i,t\text{-}k}$.

${}_{n}P_{x,i,t}$ is the population aged x to x + n in area i at the most recent census (t),

k is the number of years between the most recent censuses at time t for area i and the one preceding it for area i at time t-k.

Nuances and details of the CCR approach are provided in Swanson et al. (2016), so I do not go into them here. However, I do note here that the CCR approach can be expressed in terms of the fundamental demographic equation:

$$P_{i,t+k} = P_{i,t} + B_i - D_i + I_i - O_i \tag{3}$$

where

$P_{i,t}$ Population of area i at time t (e.g., the launch date).

$P_{i,t+k}$ Population of area i at time t + k (e.g., the projection target date).

B_i Births in area i between time t and t + k.

D_i Deaths in area i between time t and t + k.

I_i In-migrants in area i between time t and t + k.

O_i Out-migrants in area i between time t and t + k.

Equation (1) can be expressed as

$$nCCR_{x,i} = nP_{x,i,t}/nP_{x-k,i,t-k} \tag{4}$$

since

$$nCCR_{x,i} = \left(nP_{x-k,i,t-k} + B_i - D_i + I_i - O_i\right)/\left(nP_{x-k,i,t-k}\right) \tag{4.a}$$

while Eq. (2) can be expressed as

$$nP_{x+k,i,t+k} = \left(nCCR_{x,i}\right) * \left(nP_{x,i,t}\right) \tag{5}$$

since

$$nP_{x+k,i,t+k} = \left(\left(\left(nP_{x-k,i,t-k} + B_i - D_i + I_i - O_i\right)\right)/\left(nP_{x-k,i,t-k}\right)\right) * \left(nP_{x,i,t}\right) \tag{5.a}$$

where x + k > = 10 then

$$nCCR_{x,i} = \left(nP_{x-k,i,t-k} - D_i + I_i - O_i\right)/\left(nP_{x-k,i,t-k}\right) \tag{5.b}$$

and since $N_i = I_i — O_i$

$$nCCR_{x,i} = \left(nP_{x-k,i,t-k} - D_i + N_i\right)/\left(nP_{x-k,i,t-k}\right) \tag{5.c}$$

where x + k > = 10

These equations clearly reveal that the CCR method expresses the individual components of change (birth, deaths, and migration) in terms of Cohort Change Ratios. This is important because it shows that it is consistent with demographic dynamics and based in demographic theory (Swanson et al., 2023).

3 The CCR Approach to Determining a Stable Population

The CCR approach simply takes the cohort change ratios found at a current point in time and holds them constant until the population reaches stability. To determine when a population has reached stability, Swanson et al. (2016) employed the well-known "Index of Dissimilarity" as an "Index of Stability" (***S***):

$$\boldsymbol{S} = 100^{*}\left\{0.5^{*}\sum\left|\left(\mathrm{np_x}/\sum \mathrm{nP_x}\right)_{\mathrm{t+y}} - \left(\mathrm{np_x}/\sum \mathrm{nP_x}\right)_{\mathrm{t}}\right|\right\} \tag{6}$$

where

y number of years between census counts/projection cycles.

x age.

n width of the age group (in years).

t year.

S compares the relative age distribution at one point in time (t + y) with the relative age distribution at the preceding point in time (t) and measures the percentage that one distribution would have to be re-allocated to match the other. ***S*** ranges from 0 to 100; a score of zero is means that there is no allocation error, and 100 means that the maximum allocation error exists. Swanson et al. (2016) used ***S*** = 0.000000 as the definition of stability but explored four less stringent levels on the temporal path to stability. Given their work, it is evident that once a population reaches ***S*** = 0.0003, it is essentially stable. In this paper, I use ***S*** = 0.000001 as the point where stability is reached, but also discuss some of the results in terms of ***S*** = 0.0003 and ***S*** = 0.005 (quasi-stability).

It is clear that the CCR method is a "numerical" (Burdon & Faires, 2011) approach to the concept of a stable population, rather than the standard analytical one. This may not be as intellectually satisfying as analytical expressions of relationships, but the CCR method is not alone in taking a numerical approach to the examination of the stable population concept. Murphy (2021), for example, links counterfactual projections with stable population theory in examining the determinants of population aging. Moreover, in discussing the explorations by Kim and Sykes (1976) regarding stable population concepts, Cohen (1979: 286) observed that their numerical experiments uncovered empirical regularities that invited theoretical explanation. In addition to this potential benefit, it is worth noting here that the CCR method is not only based in demographic theory but consistent with Burch's (2018) perspective on demographic theory.

4 Data and Methods

A range of methods exist for estimating ***r***. However, unlike the CCR approach, they ignore migration (Coale, 1972; Keyfitz & Flieger, 1967; Lotka, 1907; McCann, 1973; Schoen, 2011; United Nations, 1968), with one notable exception (Preston, et al., 1989), which I will describe shortly. Given this, the bivariate and the multiple regression (with four predictor variables) models by

Swanson et al. (2016) appear to be the first regression models designed for this purpose in that they estimate ***r'***. While they may not be mathematically elegant, they have the potential to be highly tractable.

Because both of these models were constructed using the entire 62 country data set, I decided to construct a model in both (1) a more rigorous manner in the form of a randomly selected 20 case "training set," which could be validated using the remaining 32 cases; and (2) a leaner multiple regression model, one with only two predictor variables, the initial rate of population change (denoted by **r**) and the initial mean age of the population. The full 62 case data set is found in the Appendix Table 3, which also shows which countries were in the "training" set (Group A) and the remaining 32 countries that made up the "validation" set (Group B).

As far as I have been able to determine, with only one exception, no method, whether a closed form solution or a regression-based approach, exists for determining mean age of a population at stability from its initial mean age. The exception is provided by Preston et al. (1989) who developed two expressions (Eqs. 1 and 4 in their paper) for expressing the rate of change in a population's mean age. The first is in terms of contemporaneous rates of birth and death and the second in terms of the rated of rates of change at different ages within a population. These are elegant expressions, but laborious to implement.

Thus, with tractability in mind, as was the case with the model for estimating ***r'***, I constructed a multiple regression model with only two predictor variables, the same ones used in the multiple regression model for estimating ***r'***, namely the initial rate of population change, **r**, and the initial mean age of the population.

In developing both of the multiple regression models, I used, as noted earlier, 62 countries found in the U.S. Census Bureau's International Data Base, the same data set employed by Swanson et al. (2016) in generating their models. However, also as noted earlier, I divided it into a 30-case set used to generate the regression model and a 32-case set to validate it. The 30 cases used to generate the model were selected via the randomization algorithm found in the NCSS (release 12) statistical package (NCSS, 2023a), which I used to divide the 62 cases into the two groups, the 30 case "training" set and the 32 case "validation" set.

Using the NCSS (2023b) "basic" multiple regression procedure found in release 12, the following regression models for estimating ***r'*** and mean population age at stability, respectively, were constructed from the 30-case training data.

$$\begin{gathered} \boldsymbol{r}' = -0.01065 + (0.0000315 * \text{Initial Mean Age}) + (1.19615 * \mathbf{r}) \\ (p < 0.05)\ p < 0.05\ p < 0.05 \\ n = 30 \\ \text{adjusted } \mathrm{R}^2 = .89 \end{gathered} \tag{7}$$

$$\begin{gathered} \text{Mean Age at Stability} = 13.315 + (0.8272 * \text{Initial Mean Age}) - (192.34 * \mathbf{r}) \\ (p < 0.05)\ p < 0.05\ p < 0.05 \\ \mathrm{n} = 30 \\ \text{adjusted } \mathrm{R}^2 = .80 \end{gathered} \tag{8}$$

The multiple coefficients of determination (R^2) and the statistical significance of the parameters found in both of these models suggest that both ***r'*** and the mean age of a population on attainment of stability can be estimated for national level populations from initial conditions (**r**, the rate of population change when a population starts on the path to stability and the mean age of a population when it starts on the path to stability) with a reasonable degree of accuracy. When these models are applied to the 32-case "validation" set, they both result in a low level of error. For the ***r'*** model, the (unweighted) mean error is 0.000000185, the standard deviation of the error, 0.0053, and the maximum error is 0.0144. Given that the mean value of the estimated ***r'*** value for these same 32 countries is -0.0055989 with a standard deviation of 0.008655, these levels of error suggest that the ***r'*** model can produce reasonable estimates of ***r'*** for national level populations.

For the model that estimates mean age at stability the (unweighted) mean error is 0.12, the standard deviation is 2.86, and the maximum error is 5.67, which suggests that the "mean age at stability" model can produce reasonable estimates of mean age at stability for national level populations.

Before turning to an evaluation of their accuracy at the sub-national level in the form of a case study of Canada's provinces and territories, I briefly describe these case study data.

To generate the CCRs per the descriptions found in sections II and III, I used the 2016 and 2021 provincial and territorial census count population data by age done by Statistics Canada, which were kindly supplied to me in user-friendly form by Doug Norris (Environics Analytics). I then input these data for each province/territory into a macro-enabled excel file that moves the population in question to stability using the CCR approach, yielding, among other characteristics, the actual ***r'*** and actual mean age at stability for each province and terriitory.[1] These actual values of ***r'*** provide, respectively, the benchmarks against which the corresponding estimates produced by the national level models will be evaluated when the latter are applied to the provincial and territorial data.

5 Results

As can be seen in Table 1, Newfoundland/Labrador and the Northwest Territories both have initial negative rates of population change while the remaining provinces and territories have positive ones. Newfoundland/Labrador and the Northwest Territories maintain negative rates of change throughout the transition and in stability. During the transition to stability, they are joined in the negative change group by all but six (British Columbia, Manitoba, Ontario, Prince Edward Island, Quebec and the Yukon). The other seven switch from positive to negative rates of

[1] A detailed description of the macro that moves a population to stability is found in Swanson et al. (2016: 240). It employs a 16 × 16 Leslie Matrix in which the age-specific CCRs appear in the diagonals. The VBA code for this macro is provided in the appendix. The Province/Territory excel workbooks that use this macro along with their respective input data and results are available from the author.

population relatively early in the transition to becoming stable. This means that like Newfoundland/Labrador and the Northwest Territories, these seven provinces and territories will be stable but with a declining population. However, it will take a very long time for them to reach zero. Even for Both Newfoundland/Labrador and the Northwest Territories, which will be the first to reach zero, they will only do so around the time they attain stability. Using the definition whereby ***S*** = 0.000001, this would be approximately 480 years from now, in the year 2500. However, using the definition as discussed earlier, ***S*** = 0.0003, Newfoundland/Labrador would reach stability approximately 225 years from now and the Northwest Territories 230 years from now.

The Yukon has the highest initial rate of population change (0.021016), followed by Prince Edward Island (0.013151) and British Columbia (0.012127). All three maintain positive rates of change during the transition to stability and upon attaining stability.

In terms of the estimated rate of population change, ***r'***, Table 1 shows that the unweighted mean error (estimated – actual) is low at 0.000693. This suggests that the model (Eq. 7) has a slight tendency toward over-estimation but has a sufficient level of accuracy to be used with sub-national populations.

As was the case with the 32-case validation national-level data set, it is important to keep the context in mind when considering the accuracy of the national level ***r'*** model when it is applied to the subnational provinces and territories of Canada. First, the mean of the actual ***r'*** for Canada's 13 provinces and territories is -0.001321, the standard deviation is 0.008198, the maximum value is 0.014527 and the minimum value is -0.01836. The mean of the estimated ***r'*** for Canada's 13 provinces and territories is -0.002014, the standard deviation, 0.00854322, the maximum value, 0.015763 and the minimum value, -0.017092. As can be seen in Table 1, the largest difference between the estimated ***r'*** and the actual ***r'*** is found for British Columbia: 0.004198. This error would generate over the course of 25 years of stability approximately a 11 percent difference between the population generated by the estimated value of ***r'*** (0.0052) and the actual value of ***r'*** (0.001002; over 50 years the difference would generate a difference of approximately 23 percent. For the remaining provinces and territories, the differences are far less. Considering Nova Scotia, for example, the error between its estimated ***r'*** (-0.000583) and actual ***r'*** (-0.000918) would generate a 0.84 percent difference over the course of 25 years of stability and a 1.69 percent difference over 50 years.[2] For context, the U.S. Census Bureau found average percent errors of 2.6 percent in an evaluation of state-level forecasts done over a ten year horizon (Campbell, 2002) with a range from -7.53 percent (Nevada) to 4.71 percent (Wyoming). Given these findings over a ten year forecast horizon, a maximum error of 11 percent over a 25 year horizon is not untoward.

Table 2 provides the evaluation results of using the mean population age model in conjunction with Canada's provinces and territories. As indicated by the values

[2] These effects were found by setting 100,000 as the population at stability, applying the estimated and actual values to an exponential model ($P_t = P_0 e^{rt}$),.and generating the populations expected at selected years starting from the common radix under the estimated and actual values.

Table 1 Evaluation of the accuracy of the regression model for estimating ***Intrinsic r’*** among Canadian provinces and territories

Province/Territory	Initial rate of population change (r)	Initial mean age	Estimated Intrinsic rate of population change(r’)	Actual Intrinsic rate of population change (r’)	Difference (Estimated—Actual)
Alberta	0.006152	38.61	-0.002075	-0.002814	0.000739
British Columbia	0.012127	42.68	0.005200	0.001002	0.004198
Manitoba	0.007395	40.03	-0.000544	0.001657	-0.002201
New Brunswick	0.005157	45.38	-0.003052	-0.001595	-0.001457
Newfoundland & Labrador	-0.006419	47.94	-0.016818	-0.018360	0.001542
Northwest Territories	-0.006371	37.42	-0.017092	-0.016115	-0.000977
Nova Scotia	0.007238	44.74	-0.000583	-0.000918	0.000335
Nunavut	0.001346	28.63	-0.008138	-0.002345	-0.005793
Ontario	0.009173	42.29	0.001654	0.000514	0.001140
Prince Edward Island	0.013151	43.23	0.006442	0.008142	-0.001700
Quebec	0.005753	43.35	-0.002403	0.000290	-0.002693
Saskatchewan	0.004053	40.25	-0.004534	-0.001152	-0.003382
Yukon	0.021016	40.46	0.015763	0.014527	0.001236

Source: Statistics Canada (via Environics Analytics) with computations by author

Table 2 Evaluation of the "Mean Age" regression model for estimating mean age at stability among Canadian provinces and territories

Province/Territory	Initial rate of population change (r)	Initial mean age	Estimated mean age at stability	Actual mean age at stability	Difference (Estimated—Actual)
Alberta	0.006152	38.61	44.07	46.92	-2.85
British Columbia	0.012127	42.68	46.29	51.14	-4.85
Manitoba	0.007395	40.03	45.00	44.03	0.97
New Brunswick	0.005157	45.38	49.86	49.42	0.44
Newfoundland & Labrador	-0.006419	47.94	54.20	54.76	-0.56
Northwest Territories	-0.006371	37.42	45.49	46.16	-0.70
Nova Scotia	0.007238	44.74	48.93	49.94	-1.01
Nunavut	0.001346	28.63	36.74	30.71	6.03
Ontario	0.009173	42.29	46.53	48.26	-1.73
Prince Edward Island	0.013151	43.23	46.54	46.91	-0.37
Quebec	0.005753	43.35	48.07	47.34	0.73
Saskatchewan	0.004053	40.25	45.83	43.68	2.15
Yukon	0.021016	40.46	42.74	44.44	-1.70

Source: Statistics Canada (via Environics Analytics) with computations by author

found in Table 2, the initial mean ages are high. Across all 13 of the provinces and territories, the unweighted mean is 41 years. Nunavut has the lowest initial mean age (38.63) and Newfoundland/Labrador, the highest (47.94).

By the time stability is attained (***S*** = 0.000001), the average unweighted mean age is 46.18 years. The average unweighted mean calculated from the model's estimates is 46.36, which is slightly below to the actual (unweighted) mean. Not surprisingly, Newfoundland/Labrador will have the highest mean age at stability, 54.76 and Nunavut the lowest, 30.71.

The process of attaining stability is a long one for Canada's provinces and territories, with approximately 506 years as the average across all of the 13 provinces and territories when ***S*** = 0.000001 is used to define the attainment of stability. For New Brunswick, the province with the least time, it takes 470 years. It takes Alberta 550 years to attain stability, which is the longest. Given that generational length is approximately 30 years (Tremblay & Vézina, 2000), on average, it will take Canada's provinces and territories 17 generations on average to attain stability. However, when ***S*** = 0.0003 is used to define the attainment of stability, the average number of years to stability across all of the 13 provinces and territories is cut approximately slightly more than half, to about 240 years. Under this definition, New Brunswick reaches stability in 220 years (approximately seven generations) and Alberta, 270 years (approximately nine generations). Under this same definition, it will take eight generations on average to attain stability across all 13 provinces and territories. These findings are consistent with those by Gerland et al.(2014) that the population of the world as a whole is unlikely to reach stability in this century.

Although thinking in terms of 250 or so years and eight generations into the future may not be as abstract as thinking in terms of 500 years and 17 generations, even this this time frame seems to be out of our reach. However, not so out of reach is the likelihood that within this century, most if not all of the provinces as well as Canada as a whole will achieve "quasi-stability." That is, when $\boldsymbol{S}$ reaches 0.005. Evidence for this statement can be seen in Figs. 1 and 2, which suggest that within 100 years, Alberta and Nunavut, respectively, will reach $\boldsymbol{S}=0.005$ if current CCRs remain constant. In turn, this suggests that quasi-stability would be achieved in less than three generations and also would be achieved within the lifetimes of Canadians born since 2020, given that life expectancy at birth for both sexes remains around 81.55 years (Statistics Canada, 2023). This is likely to occur because the timeframe for reaching $\boldsymbol{S}=0.005$ shown in Figs. 1 and 2 for these two examples are virtually the same for the rest of Canada's provinces and territories.

The path to stability is not always a smooth one. During, the initial stage of the transition to stability, the Stability Index typically does not decrease monotonically. It takes a while for the initial age groups to be "ironed out" by the application of the constant CCRs in regard to most of the provincial and territorial populations. Figure 1 provides an idea of this process using Alberta as an example, which applies to most of the provinces/territories. Here, one can see that it takes about 150 years before the path becomes monotonically smooth, which is about five generations. An exception in this regard is Nunavut, which displays a path during the transition whereby the stability index decreases monotonically and smoothly. This can be seen in Fig. 2.

6 Discussion

In evaluating the two regression models based on national-level data that provide sub-national estimates, respectively, of ***r'*** and mean age at stability, I find them to be sufficiently accurate to be useful. However, as with the case with many (if not most) regression models, there are "outliers" in terms of the differences between the estimated and actual values. This is particularly important to keep in mind in terms of the ***r'*** model and far less important in terms of the mean age model. Given this caveat, the models are useful because they provide information not available from the traditional analytical approaches. As such, these results support the use of: (1) The Cohort Change Ratio (CCR) approach in examining stable population concepts; and (2) the two multiple regression models for estimating ***r'*** and the mean age of a population at stability that are built around this approach.

Evaluating the subnational accuracy of models constructed using national level data is useful because it is often the case that in many countries the age data needed to construct CCRs at the subnational level may not be as available as data at the national level. Canada's provinces and territories were selected for the subnational accuracy not only because of availability but also because of their high quality.

These findings also show that there are links between factors found during initial conditions and their counterparts when a population attains stability. In turn, they

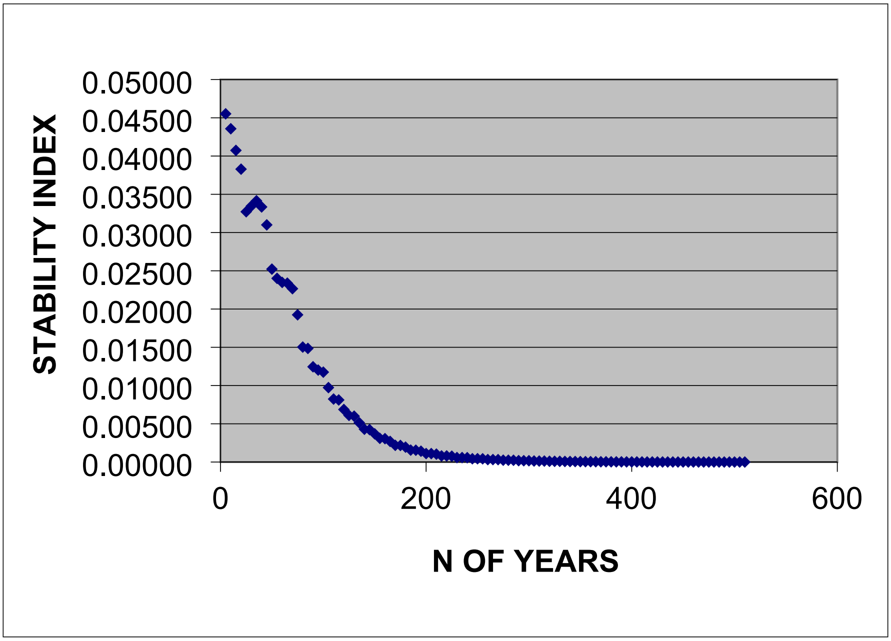

Fig. 1 Alberta: Path to stability

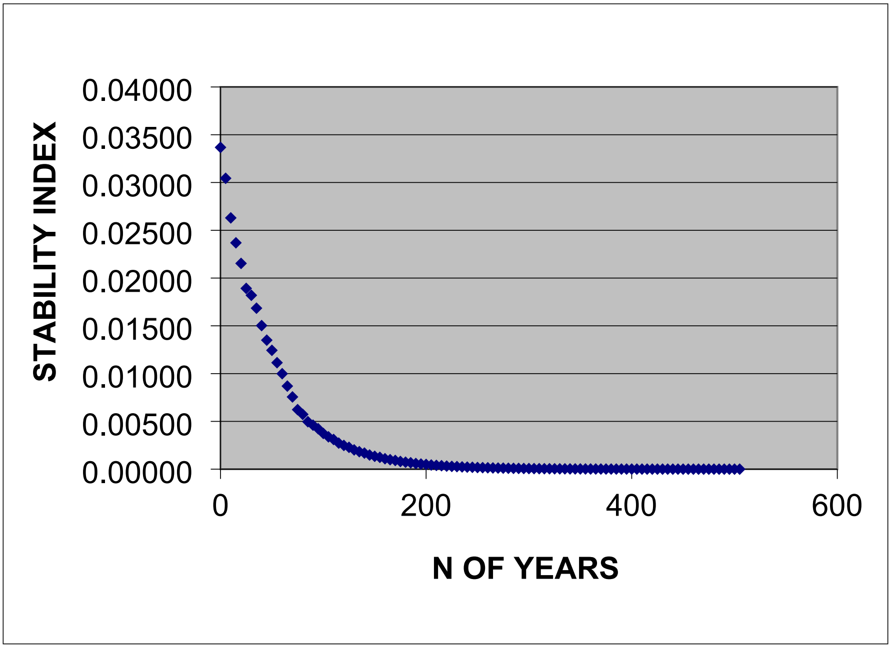

Fig. 2 Nunavut: Path to stability

suggest that there are overlooked factor linkages because the widespread knowledge and acceptance of the ergodic nature of the "age structure factor" may be "masking" the possibility that ergodicity is not universal in regard to other factors. That is, while ergodicity applies to a population's initial age structure, it does not always apply to other initial conditions, even those associated with initial age structure – mean age, as this paper shows, is notable in this regard. This possibility is consistent with work by both Wachter (1986) and Stott et al. (2010: 242), the latter of whom find that "reducible" models (those in which parameterized transition rates do not facilitate pathways from all stages to all other stages) based on population projection matrices are sometimes ergodic but may be non-ergodic (the model exhibits two or more stable asymptotic states with different asymptotic stable growth rates, which depend on the initial stage structure used in the population projection).

I conclude by noting that regression models are generally not as intellectually satisfying as analytical expressions in regard to describing relationships. However, in discussing the explorations of stable populations by Kim and Sykes (1976), Cohen (1979: 286) observes that their numerical experiments uncovered empirical regularities that invited theoretical explanation. In a similar vein, it is worth recalling that prior to the introduction of the CCR-based models, regression analysis had already been successfully applied in conjunction with stable population analysis. These applications include the Bourgeois-Pichat method for estimating $\boldsymbol{r}$ from the proportional age distribution of a given population (Keyfitz & Flieger, 1967 168:49, United Nations, 1968), McCann's (1973) method for estimating mean generation length from a trial value of the intrinsic rate of increase, and the generation of model life table families and from them, stable populations (Coale & Demeny, 1966). Along with the applications presented here, they suggest that there may be more.

In defense of non-analytic methods, Daoud and Dubhashi (2023: 32), following the insights of Breiman (2001), point out that predictive statements (as found in CCR approach to stable population theory discussed in this paper) and causal inference (as found in the analytic approach to stable population theory) fall along the fault line between two cultures of statistical modeling, the algorithmic modeling culture (AMC) and the data modeling culture (DMC). They go on to identify the hybrid modeling culture (HMC), an emerging approach that has evolved and mutated from AMC and DMC. While the CCR approach falls within AMC, this paper shows that it can be used in model validation, a necessary component of the DMC approach. As such, it represents a combination of the two statistical modeling cultures.

Given our limited knowledge of the factors that correspond between initial conditions and the attainment of stability, further explorations of these linkages appear to be in order. Whether explorations based on the CCR approach lead to a "paradigm shift" or formal demography remains in its current "normal science" phase is an open question, but either way, as Kuhn (1962: 15) observes: "History suggests that the road to a firm research consensus is extraordinarily arduous."

Appendix

3

Table 3 Input data for the regression models

Country	Initial Mean Age	Mean Age at Stability	Initial r	R at stability (r')	Randomly Assigned Group
Albania	33.46	41.79	-0.009775	-0.01558	A
Armenia	33.70	44.18	-0.000675	-0.013549	B
Australia	38.09	43.88	0.011112	0.003347	A
Austria	41.48	45.84	0.002042	-0.004820	B
Azerbaijan	30.38	38.76	0.009732	-0.000681	B
Belarus	39.11	47.14	-0.004487	-0.016756	B
Belgium	41.34	45.03	0.001154	-0.004502	B
Bosnia-Herzegovina	37.25	48.40	0.003118	-0.012648	B
Bulgaria	41.66	47.04	-0.009828	-0.019276	A
Canada	39.92	42.11	0.006990	-0.002409	A
Chile	32.42	41.82	0.009508	-0.001179	A
Costa Rica	29.18	38.99	0.015827	0.004282	B
Croatia	40.57	46.12	0.002966	-0.004958	B
Cuba	35.88	45.35	0.000892	-0.011471	A
Czech Republic	40.34	49.32	-0.000379	-0.014536	B
Denmark	40.10	41.43	0.002757	-0.000069	B
El Salvador	26.52	34.32	0.003317	-0.004135	B
Estonia	40.41	46.73	-0.007492	-0.016026	A
Fiji	27.97	33.32	0.008792	0.000942	B
Finland	40.97	49.32	-0.000670	-0.011473	A
France	39.93	43.37	0.005009	-0.000288	A

Table 3 (continued)

Country	Initial Mean Age	Mean Age at Stability	Initial r	R at stability (r')	Randomly Assigned Group
Georgia	38.14	46.13	-0.001652	-0.011565	B
Germany	42.83	48.02	-0.000402	-0.008295	B
Greece	42.25	50.36	0.001026	0.011066	A
Guatemala	23.11	25.83	0.018257	0.016679	A
Hong Kong	40.01	59.32	0.005890	-0.017350	A
Hungary	40.48	46.98	-0.002140	-0.012480	A
Ireland	35.69	40.74	0.021030	0.012600	A
Israel	32.05	35.32	0.018770	0.014630	A
Italy	43.41	50.14	0.003110	-0.007350	A
Jamaica	27.64	31.89	0.008486	0.004457	A
Japan	43.81	53.89	0.000020	-0.12752	A
Kazakhstan	31.07	36.58	0.006000	-0.001652	B
Kyrgyzstan	27.70	32.23	0.010240	0.004827	B
Latvia	40.35	47.43	-0.007540	-0.018482	A
Lithuania	39.47	50.66	-0.003380	-0.017554	B
Macedonia	35.55	41.88	0.003370	-0.006019	B
Moldova	35.51	44.13	-0.011140	-0.023239	B
Montenegro	37.32	51.88	-0.007150	-0.025181	B
Netherlands	39.53	43.01	0.003340	-0.002198	B
New Zealand	36.89	41.04	0.011596	0.006334	B
Norway	39.36	42.02	0.005530	0.001347	A
Poland	38.48	49.45	-0.000510	-0.015138	A
Portugal	40.44	47.07	0.003775	-0.005835	A
Romania	38.81	47.81	-0.002732	-0.015953	B

Table 3 (continued)

Country	Initial Mean Age	Mean Age at Stability	Initial r	R at stability (r')	Randomly Assigned Group
Russian Federation	38.77	45.54	-0.005484	-0.016634	B
Saudi Arabia	24.68	27.28	0.025692	0.019951	B
Serbia	40.81	44.72	-0.003917	-0.010320	B
Singapore	34.04	41.47	0.023416	0.012097	A
Slovakia	37.52	48.25	0.000765	-0.014512	A
Slovenia	40.93	50.27	-0.000708	-0.014502	B
Spain	41.04	48.19	0.013702	0.002383	A
Sweden	41.54	43.93	0.003166	-0.000912	B
Switzerland	40.96	45.24	0.003738	-0.002453	A
Tajikistan	24.21	28.55	0.019027	0.013821	A
Turkmenistan	26.41	33.08	0.013357	0.005719	B
Ukraine	40.08	47.62	-0.008420	-0.020649	B
United Kingdom	40.06	42.05	0.004786	0.001357	A
Uruguay	35.64	38.73	0.002286	-0.001422	B
USA	37.37	40.13	0.008780	0.004692	A
Uzbekistan	26.75	34.91	0.013780	0.007176	A
Venezuela	27.51	33.46	0.013780	0.007176	B

Source: International Data Base, U.S. Census Bureau (see Swanson et al., 2016)

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s42650-024-00080-6.

Acknowledgements I am grateful to: Doug Norris (Environics Analytics | environicsanalytics.com) for kindly providing the 2016 and 2021 provincial and territorial population data by age in a "user-friendly" format; Alison Yacyshyn (Concordia University of Edmonton) for reviewing an earlier draft and providing helpful comments, the editor of CSIP for suggestions; and the comments made by two anonymous reviewers.

Funding This research was not funded by any agency.

Data Availability The data underlying this paper are secondary and available from the author.

Declarations A clinical trial registration statement is not applicable because no human subjects were involved.

A statement regarding permission to reproduce material from other sources is not applicable because no excerpts from copyrighted works owned by third parties are included.

Ethical Approval An ethics approval statement is not applicable because the data used are secondary and no human subjects review was required.

Patient Consent A patient consent statement is not applicable.

Conflict of Interest The author has no conflicting interests in regard to this paper.

References

Arthur, W. B. (1982). The ergodic theorems of demography: A simple proof. *Demography, 19*(4), 439–445.

Breiman, L. (2001). Statistical modeling: The two cultures. *Statistical Science, 16*(3), 199–231.

Burch, T. (2018). *Model-Based Demography: Essays on integrating data, technique, and theory*. Demographic Research Monographs. Rostock, Germany. Max Planck Institute for Demographic Research.

Burdon, R., & Faires, J. (2011). *Numerical Analysis* (9th ed.). Brooks/Cole.

Campbell, P. (2002). Evaluating forecast error in state level population projections using Census 2000 counts. Working paper POP-WP057. U.S. Census Bureau. https://www.census.gov/library/working-papers/2002/demo/POP-twps0057.html. Accessed 1 Aug 2023

Carmichael, G. (2016). *Fundamentals of demographic analysis: Concepts*. Springer. B.V. Press.

Caswell, H. (2001). *Matrix population models: Construction, analysis, and interpretation* (2nd ed.). Sinauer Associates Inc.

Coale, A. J. (1972). *The growth and structure of human populations: a mathematical investigation*. Princeton University Press.

Coale, A. J., & Demeny, P. (1966). *Regional model life tables and stable populations*. Princeton University Press.

Cohen, J. (1979). Ergodic theorems in demography. *Bulletin of the American Mathematical Society, 1*(2), 275–295.

Daoud, A., & Dubhashi, D. (2023). Statistical modeling: The three cultures. *Harvard Data Science Review,* 5.1. https://hdsr.mitpress.mit.edu/pub/uo4hjcx6/release/1. Accessed 5 Aug 2023
Dublin, L., & Lotka, A. (1925). On the true rate of natural increase: As exemplified by the population of the United States, 1920. *Journal of the American Statistical Association, 20*, 305–339.
Gerland, P., et al. (2014). World population stabilization unlikely this century. *Science, 346*(62006), 234–237.
Keyfitz, N., & Flieger, W. (1967). *World population: An analysis of vital data*. University of Chicago Press.
Kim, Y., & Sykes, Z. (1976). An experimental study of weak ergodicity in human populations. *Theoretical Population Biology, 10*, 150–172.
Kuhn, T. (1962). *The structure of scientific revolutions*. University of Chicago Press.
Lopez, A. (1961). *Problems in stable population theory*. Office of Population Research. Princeton University.
Lotka, A. (1907). Relation between birth rates and death rates. *Science (New Series), 26*(653), 21–22.
McCann, J. (1973). A more accurate short method of approximating Lotka's r. *Demography, 10*(4), 567–570.
Murphy, M. (2021). Use of counterfactual population projections for assessing the demographic determinants of population ageing. *European Journal of Population, 37*(1), 211–242.
NCSS. (2023a). Randomization lists. https://www.ncss.com/wp-content/themes/ncss/pdf/Procedures/NCSS/Randomization_Lists.pdf. Accessed 15 Jul 2023
NCSS. (2023b). Basic multiple regression. https://www.ncss.com/wp-content/themes/ncss/pdf/Procedures/NCSS/Multiple_Regression-Basic.pdf. Accessed 15 Jul 2023
Popoff, C., & Judson, D. (2004). Some methods of estimation for statistically underdeveloped areas. In J. Siegel & D. A. Swanson (Eds.), *The methods and materials of demography* (2nd ed., pp. 603–641). Elsevier Academic Press.
Preston, S., Himes, C., & Eggers, M. (1989). Demographic conditions responsible for population aging. *Demography, 26*(4), 691–704.
Preston, S., Heuveline, P., & Guillot, M. (2001). *Demography: Measuring and modeling population processes*. Blackwell Publishing.
Schoen, R. (2011). Age-specific growth, reproductive values, and intrinsic r. *Demographic Research, 24*. http://www.demographic-research.org/Volumes/Vol24/33/. Accessed 1 Feb 2024
Sharpe, F., & Lotka, A. (1911). A problem in age distribution. *Philosophical Magazine, Series, 1*(2), 425–438.
Statistics Canada. (2023). Canada, Both Sexes, Life expectancy (in years) at age x (ex), 2020–2022. https://www150.statcan.gc.ca/t1/tbl1/en/tv.action?pid=1310011401. Accessed 10 Jan 2024
Stott, I., Townley, S., Carslake, D., & Hodgson, D. (2010). On reducibility and ergodicity of population projection models. *Methods in Ecology and Evolution, 1*, 242–252.
Swanson, D. A., Tedrow, L., & Baker, J. (2016). Exploring stable population concepts from the perspective of cohort change ratios: Estimating the time to stability and intrinsic *r* from initial information and components of change". In R. Schoen (Ed.), *Dynamic demographic analysis* (pp. 227–258). Springer B.V. Press.
Swanson, D. A., Bryan, T., Hattendorf, M., Comstock, K., Starosta, L., & Schmidt, R. (2023). An example of combining expert judgment & small area projection methods: Forecasting for water district needs. *Spatial Demography*. https://link.springer.com/article/10.1007/s40980-023-00119-3. Accessed 5 Jan 2024
Tremblay, M., & Vézina, H. (2000). New estimates of intergenerational time intervals for the calculation of age and origins of mutations. *American Journal of Human Genetics, 66*(2), 651–658.
Tuljapurkar, S. (1982). Population dynamics in variable environments. IV: Weak ergodicity in the Lotka equation. *Journal of Mathematical Biology, 14*, 221–230.
United Nations. (1968). *The concept of a stable population: applications to the study of populations with incomplete demographic statistics*. Population Studies No. 39, Department of economic and social affairs. United Nations.
United Nations. (2017). World population ageing 2017 highlights. New York, NY. Department of economic and social affairs. United Nations. https://www.un.org/en/development/desa/population/publications/pdf/ageing/WPA2017_Highlights.pdf. Accessed 10 Aug 2024
United Nations. (2002). *Methods for estimating adult mortality*. Population Division, United Nations.
Wachter, K. (1986). Ergodicity and inverse projection. Population Studies 40(2), 275–287. https://www.tandfonline.com/doi/abs/10.1080/0032472031000142076. Accessed 25 Feb 2024

Yusuf, F., Martins, J., & Swanson, D. A. (2014). *Methods of demographic analysis*. Springer B.V. Press.